Reviewer’s report

Title: Variations in cardiovascular disease under-diagnosis in England: national cross-sectional spatial analysis

Version: 1 Date: 16 September 2010

Reviewer: John Robson

Reviewer’s report:

Major Recommendation.

The authors should provide more detail on assumptions made about constructing estimate population age structure and associated demographic characteristics in both the GP and HSSE samples in both their summary and methods. In their discussion they should acknowledge the limitations of these data and their assumptions both about population demography and smoking. They should seriously consider the possibility that their findings, particularly in areas associated with high ethnic density, might result from model error or artefact.

Summary:

The authors should make it clear in the summary that the expected prevalence is based upon patient reports of disease and the observed prevalence is based upon doctor recorded diagnoses in general practice and that the denominator populations used for the observed prevalence in general practices are based upon the assumption that their demographic structure is the same as that of the wider Local Authority area within which they are located.

Methods:

“The core model variables are (ten-year age band), gender (male and female), ethnicity (Asian/Asian British, Black/Black British and White) and deprivation (quintiles of Index of Multiple Deprivation 2004 scores)21.” For CHD smoking is added. The authors should give details of the coefficients/ risk ratios associated with the variables in these models in particular the ratios/coefficients associated with ethnicity and smoking are of interest.

Several assumptions would benefit from further clarification, with clear separation in the text between the assumptions made about the denominator population structures and assumptions made about risk factors.

Denominator populations: It should be made clear that QoF data do not provide the age structure of the populations from which they are derived and cannot be age-standardised.

The authors ‘Matched practice populations to LA areas’ – but practices do not ‘match’ to LA areas. The assumption that a practice age structure and
demography is the same as the entire Local Authority within which it is located may introduce substantial error and ecological confounding. Practices within a single LA area may differ substantially in their demographic characteristics and may cross LA boundaries.

Rivas C, Bobby J, Boomla K, Robson J. Hospital data may be more accurate than census data in estimating the ethnic composition of general practice populations. Inform Prim Care. 2009;17(2):67-78.


The authors assume “that the distribution of ethnic populations is uniform across quinary age bands.” This is not the case as many immigrant populations have the highest birthrates in the UK and their population structure is notably younger than White people. It should be made clear that this assumption will introduce error. Furthermore, disease prevalence in ethnic groups within the HSfE is based upon a ‘boost’ sample (as the national sample contained too few people from ethnic minorities). The authors give no detail as to how this boost sample was related to the ethnic distribution of the population denominators within the broader national sample.

Explanatory variables: Assumptions about explanatory variables in the model should be considered in more detail.

Under limitations it should be stated that the model does not contain major risk factors known to be predictive of vascular disease including blood pressure. In addition smoking data are almost 10 years out of date and “Model assumptions include that the proportion of smokers, ex-smokers and never-smokers is uniform across ethnic categories ...” . However White populations have far higher rates of smoking than Black populations and this assumption is incorrect.

The authors ask “An obvious question is: why was there such a discrepancy between expected and practice registered disease prevalence, when the model-based CHD and stroke prevalence estimates were based on patient reports of doctor-diagnosed disease?”

They consider the most likely reason is underdiagnosis. This seems very unlikely as it would require substantial differences in organisational structure and behaviour between practices in neighbouring or socially comparable geographical areas. The authors offer three reasons why this may be the case but provide no evidence to support any of them. Inadequate searches and coding in GP records, failure to record hospital diagnoses and population mobility. Inadequate searching or failure to record hospital diagnoses is an unlikely reason because validating diagnoses on CHD registers for the Quality and Outcomes Framework has led to substantial improvement and uniformity in recording throughout England with no systematic geographical differences.

Carey IM, Dewilde S, Harris T, Whincup PH, Cook DG. Spurious trends in coronary heart disease incidence: unintended consequences of the new GP
Population mobility is an unlikely explanation for area differences. Population mobility is greatest at younger ages who do not have vascular disease. Mobility is less pronounced at older ages and does not explain differences within mobile urban areas like London or Leicester which display extremes of discrepancy in both directions.

The authors have given insufficient consideration to the more likely possibility that their findings may result from artefact introduced by erroneous assumptions in the model. These assumptions relate both to the denominator populations used and to the explanatory variables, particular in relation to ethnic distributions of the populations and smoking. The denominator populations are based upon assumptions of equivalence between practice and local authority populations in regard to age and ethnic composition which may be further compounded by assumptions for the HSE ethnic ‘boost’ sample.

The limited information on standard vascular risk factors including blood pressure and cholesterol used as essential components of CHD risk prediction models like Framingham is a major omission. The only behavioural risk factor included is smoking more than a decade old and likely to be erroneous for ethnic groups.

There is an alternative explanation for discrepancy between observed and expected differences, namely a poorly performing model.


This is illustrated by Figure 1 which shows that the greatest discrepancy/instability (both over and under-estimation) between observed and expected rates for CHD tend to occur in spatial association with areas with the highest ethnic density. The authors should consider the conclusion that this model performs poorly in areas with high ethnic density because of the limitations behind several assumptions upon which it is based.

Review by a statistician with an interest in this field would be of assistance: such as Michael Goldacre

Professor of Health Care Epidemiology: Department of Public Health, University of Oxford, Old Road Campus, Headington, Oxford OX3 7LF, UK

Level of interest: An article whose findings are important to those with closely related research interests
Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

'I declare that I have no competing interests'